JOURNAL OF APPLIED BEHAVIOR ANALYSIS 2008, **41,** 417–422 NUMBER 3 (FALL 2008)

USING PERFORMANCE FEEDBACK TO IMPROVE TREATMENT INTEGRITY OF CLASSWIDE BEHAVIOR PLANS: AN INVESTIGATION OF OBSERVER REACTIVITY

ROBIN S. CODDING

CITY UNIVERSITY OF NEW YORK
GRADUATE SCHOOL AND UNIVERSITY CENTER

ANDREW LIVANIS

LONG ISLAND UNIVERSITY

GARY M. PACE

MAY INSTITUTE

AND

LESLIE VACA

LONG ISLAND UNIVERSITY

The current study replicated the positive effects of performance feedback on treatment integrity and extended previous work by examining reactivity using a multiple baseline design with alternating treatments for observer-present and observer-absent conditions on teachers' implementation of a classwide behavior plan. No differences were found between conditions, and treatment integrity improved across all teachers, suggesting that performance feedback, rather than observer reactivity, was responsible for reported behavior changes.

DESCRIPTORS: observer reactivity, performance feedback, treatment integrity

The demonstration of functional relations between independent and dependent variables is central to applied behavior analysis. However without clear evidence that an independent variable is implemented as intended, there can be no definitive conclusions regarding a functional analysis of behavior (Johnston & Pennypacker, 1980). The degree to which an independent variable is implemented as intended has come to be known as treatment integrity (Gresham, Gansle, & Noell, 1993). One successful intervention strategy for enhancing treatment integrity is performance feedback (Noell et al., 2005).

Address correspondence to Robin S. Codding, who is now at the Department of Counseling & School Psychology, College of Education, University of Massachusetts–Boston, 100 Morrissey Boulevard, Boston, Massachusetts 02125 (e-mail: Robin.codding@umb.edu).

doi: 10.1901/jaba.2008.41-417

Research has shown that performance feedback can effectively improve treatment integrity for academic (Noell et al., 2000) and behavioral problems treated with single- (Noell, Duhon, Gatti, & Connell, 2002) and multiple-component plans (Codding, Feinberg, Dunn, & Pace, 2005). For example, Codding et al. conducted treatment integrity observations on implementation of the individualized behavior support plans of 5 students and provided immediate verbal feedback to the 5 teachers who worked with these students. Effects of performance feedback provided in this manner were evaluated separately across antecedent and consequent components using a multiple baseline design across 5 teacher–student dyads. Results suggested that performance feedback, which consumed a mean of 12 min of consultants' time, led to more accurate implementation of the antecedent components for 4 of 5 teachers

and for the consequence components for 5 of 5 teachers.

Despite its documented success, it is possible that the treatment effects associated with performance feedback are confounded by reactivity to being observed (Gresham, MacMillian, Beebe-Frankenberger, & Bocian, 2000). That is, it has been suggested that those who are being observed improve their treatment integrity because of the presence of an observer rather than the effects of performance feedback. The purpose of the present study was to investigate the role of observer reactivity when using a direct observation method of assessment for treatment integrity while conducting a systematic replication of Codding et al. (2005). Specifically, performance feedback was used with teachers who were implementing a classwide behavior plan for children who had been diagnosed as emotionally disturbed and who were receiving special education services in a public school setting to determine if improvements in teachers' implementation of behavior plans occurred in the presence or absence of an observer.

METHOD

Participants and Setting

We conducted the present study in a self-contained program housed within a public general education school that consisted of two classes. The classroom participating in this study consisted of 6 seventh graders (5 boys, 1 girl) and 1 female eighth grader. The students ranged in age from 12 to 14 years. Students' diagnoses included attention deficit hyperactivity disorder (5 students), bipolar disorder (3 students), conduct disorder (3 students), and anxiety disorder (1 student).

Three teachers participated in this study. Mr. Rivers, the primary teacher, held a master's degree in special education and reading and had 3 years (2 in this program) of teaching experience. Mrs. Smith, one of the classroom aides, had 9 years (3 in this program) of experience. Miss Hill, the second aide, was hired just prior to her participation in the study and had completed a 4-year undergraduate degree. She had no teaching experience. This study was approved by two human subject institutional review boards.

The classes were located on either side of the school psychologist's office, which contained one-way observation windows and audio monitors for each classroom. Therefore, the observer could enter the office without being seen by the teachers who participated in the study.

Materials

Classroom behavior management plan. A programwide behavior intervention plan had been created 3 years prior to the start of this study and was ongoing at the time of investigation. The components of the plan consisted of specific behaviors that teachers were instructed to engage in as a result of the activity a student was required to perform, in response to a student's behavior, or at certain points of the day. This plan consisted of three general procedures: (a) a level system in which students earned points daily and successively obtained more rewards for engaging in prosocial behaviors, (b) neutral statements teachers were required to make when students were not complying with requests, and (c) time out.

Response measurement sheet. We adapted the data sheet from Codding et al. (2005) for the classroom behavior management plan, which was divided into 14 components. Each component was rated on one of three levels of implementation integrity: (a) implemented as written (i.e., the entire component was implemented every time the operational definition designated the teacher should act), (b) not implemented as written (i.e., the teacher either did not consistently implement the component every time he or she was required to, or the teacher did not implement the component at all), and (c) no opportunity to observe (i.e., there was no need for the teacher to implement the component during the period of time he or

she was being observed). Percentage of correct implementation was used as the measure of treatment integrity and was calculated by dividing the number of plan components implemented as written by the total number of plan components teachers had the opportunity to exhibit. The mean percentage of plan steps that teachers had the opportunity to exhibit during sessions was 57% (range, 38% to 86%).

Social validity. Codding et al. (2005) designed the 10-item questionnaire we administered to teachers after the performance feedback phase was terminated in order to examine the acceptability of integrity observations. Ratings ranged from 1 (*strongly disagree*) to 5 (*strongly agree*). Teacher ratings on the questionnaire were provided anonymously.

Training

Each year, the school district hired a psychologist from a local psychiatric center to provide two 4-hour training sessions for program staff. During these sessions, the psychologist reviewed and modeled aspects of the classwide plan and provided opportunities to teachers for role playing. Prior to the start and continuing throughout the duration of this study, the teachers received ongoing weekly consultation from staff at the psychiatric center that was designed to address systemic areas of concern (e.g., scheduling for mainstream classes) or overly problematic behaviors in the classroom.

Procedure

Observation sessions. Observation data were collected from behind the one-way window by an investigator two or three times per week per teacher for approximately 4 months. Each observation session during baseline and treatment lasted 44 min (the length of a school period). The observer was located inside the classroom during 50% of the observations (observer present) and was absent from the classroom (but observing from the one-way window) for the other 50% of observations (observer absent). We selected the days, times, and the conditions (observer present vs. absent) of the observations randomly to reduce the chance that the teachers would become aware of the days, times, or condition under which observations would occur (Gresham et al., 1993).

Phases. Baseline consisted of observing each teacher and completing the integrity data sheet under both observer-present and observer-absent conditions. No feedback was provided during the baseline phase, and teachers were not privy to the notes or completed response measurement sheets. Performance feedback was implemented after stable performance in baseline was demonstrated for both conditions (i.e., observer present and absent). The observer, who had previously worked with the program and school for 3 years, attempted to provide the feedback within 5 to 10 min after every observation was completed, although the observer provided some feedback after a delay of 30 to 60 min due to school constraints (e.g., lunch schedules, fire alarm drills). The observer provided feedback on all of the steps observed and included praise for steps followed as written and corrective feedback for those steps that were followed sometimes or not at all. The observer conducted the first feedback session following the last baseline session. The mean length of the feedback meeting was 10 min (range, 5 to 20 min).

Experimental design. A multiple baseline design across staff members with alternating treatments was selected to evaluate the effectiveness of performance feedback across observer-present and observer-absent conditions.

Interobserver agreement and procedural integrity. The investigator and a graduate student recorded teachers' implementation of the behavior support plan simultaneously but independently from behind the one-way window on 34% of the observation sessions. On a component-by-component basis, comparisons were

conducted between the two sets of' responses. Agreement was calculated by dividing the number of agreements per plan component by the number of agreements per plan component plus disagreements per plan component and converting this ratio to a percentage. Mean agreement was 96% (range, 79% to 100%). Procedural integrity of performance feedback meetings between each teacher and the observer was assessed by a graduate student who served as a second independent observer of the meeting, across 45% of sessions. The graduate student attended feedback meetings with the observer and recorded whether the observer followed the three-step performance feedback procedure (defined by Codding et al., 2005) that included the following: (a) attempted to provide feedback within 5 to 10 min of the observation, (b) praised plan components implemented as written, and (c) reviewed and provided constructive feedback for the components not implemented as written. Procedural integrity of performance feedback was 100%.

RESULTS AND DISCUSSION

Results of the effects of performance feedback and observer presence for Mr. Rivers, Mrs. Smith, and Miss Hill are illustrated in Figure 1. Consistent with previous research (e.g., Codding et al., 2005), performance feedback led to improvements in correct implementation that reached 100% for all teachers. During baseline, correct implementation across teachers ranged from 0% to 57%; during the performance feedback phase, implementation ranged from 50% to 100% correct. For Mr. Rivers and Miss Hill a level change was observed, whereas for Mrs. Smith overlapping data points occurred initially between conditions. Thus, in the present study we replicated the findings from Codding et al., suggesting that performance feedback led to increases in implementation accuracy for a multiple-component behavior plan. In addition, these results extended the application to a classwide behavior plan for a self-contained program within a general education school. This is an important expansion, because less training and more variability in service delivery were associated with this public education setting than in the previous study. Codding et al.'s study was conducted in a residential special education school whose programming, training, and service delivery adhered to the tenets of applied behavior analysis. This research contributes to a growing literature that has examined effective follow-up procedures that enhance behavioral consultation and directly address treatment integrity (Noell et al., 2005). To our knowledge, research applying this strategy to improve implementation of a classroom plan has not been conducted and is important for accountability as public schools move towards implementing school- and classwide behavior plans in efforts to enhance school safety and climate.

Concurring with findings from Codding et al. (2005) and Noell et al. (2002), mean ratings across teachers suggested favorable views of performance feedback (M = 4.6, 4.7, and 5.0 for Mr. Rivers, Mrs. Smith, and Miss Hill, respectively) on items related to the purpose of, procedures for, and benefits on their skills and the subsequent impact on students' behavior. Anecdotally, the teachers reported that they enjoyed receiving positive feedback on correctly implemented components. In fact, the teachers requested to administrators that external consultation from the psychiatric hospital be terminated in favor of receiving performance feedback.

Most notable in the present study was that no differentiation was observed between observer-present or observer-absent conditions during baseline or performance feedback phases. These findings suggest that observer reactivity did not contribute to the positive effects of performance feedback when using direct observation as the method of assessment. For example, mean baseline accuracy for Mr. Rivers was 23% and 21% for observer-absent and observer-present

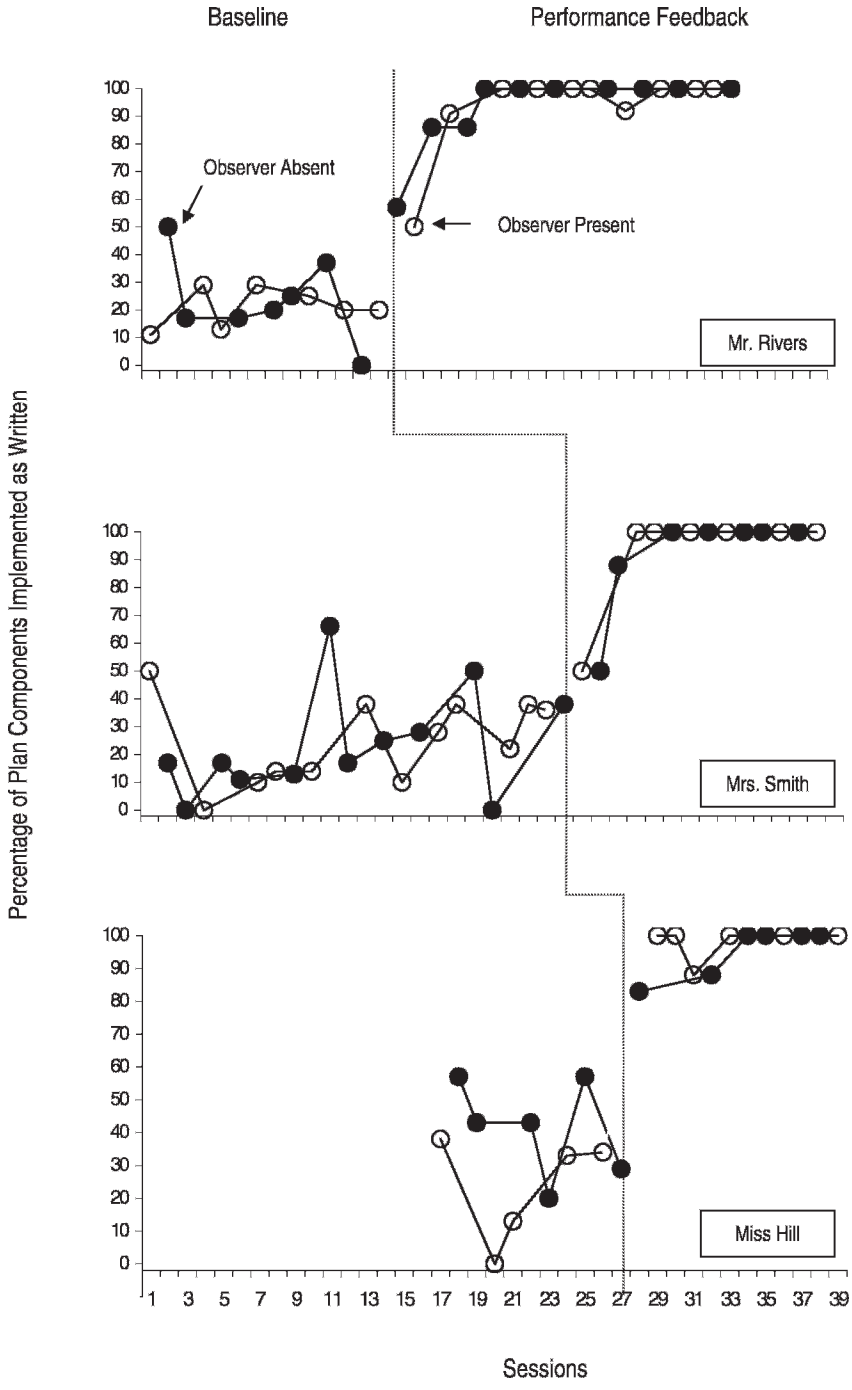

Figure 1. Percentage of classwide behavior plan components implemented as written across teachers.

conditions, respectively. During feedback, both observer conditions resulted in a mean performance of 93%. Similar findings were obtained for Mrs. Smith and Miss Hill. Therefore, performance feedback, not observer reactivity, resulted in improved treatment integrity. The results of the present study are consistent with previous research investigating observer reactiv-

ity. Johnson and Bolstad (1975) compared the behavior of family members at home using audiotape recordings (observer absent) and observer-present conditions and found no differentiation between conditions. In a study investigating teacher and student behavior using direct observations, Hay, Nelson, and Hay (1980) found that in only one of four classrooms studied did observer reactivity affect the behavior of those being observed.

Our study has several limitations that should be noted. First, the observer-present and observer-absent conditions were not independent. Therefore it is possible that the change in behavior in the observer-absent condition could be the result of a carryover effect from the observer-present condition. However, this explanation does not diminish the finding that even when the observer was not in the classroom, teacher performance improved. A second limitation is that neither the teachers nor the observer were blind to the purpose of the study. That is, they knew their behaviors were going to be observed, which may have contributed to the lack of differentiation at baseline. However, an attempt was made to reduce teachers' awareness of the days of the week and times of the day during which these observations would occur by randomly determining these sessions. Another effort to decrease teachers' knowledge that the investigator was observing them during observer-absent conditions was to enter the observation room through another classroom. Finally, student behavior was not observed simultaneously with teacher behavior to examine the correspondence between increases in correct plan implementation.

Given the repeated evidence suggesting that performance feedback is an important factor for improving plan implementation (e.g., Noell et al., 2005), future research should continue to investigate the assessment procedures used to collect treatment integrity data in various settings and across situations. Emphasis should be placed on providing evidence of reliable and accurate tools that also are efficient for school professionals.

REFERENCES

Codding, R. S., Feinberg, A. B., Dunn, E., & Pace, G. M. (2005). Effects of immediate performance feedback on implementation of behavior support plans. *Journal of Applied Behavior Analysis*, *38*, 205–219.

Gresham, F. M., Gansle, K. A., & Noell, G. H. (1993). Treatment integrity in applied behavior analysis with children. *Journal of Applied Behavior Analysis*, *26*, 257–263.

Gresham, F. M., MacMillan, D. L., Beebe-Frankenberger, M. E., & Bocian, K. M. (2000). Treatment integrity in learning disabilities intervention research: Do we really know how treatments are implemented? *Learning Disabilities Research and Practice*, *15*, 198–205.

Hay, L. R., Nelson, R. O., & Hay, W. M. (1980). Methodological problems in the use of participant observers. *Journal of Applied Behavior Analysis*, *13*, 501–504.

Johnson, W. M., & Bolstad, E. D. (1975). Reactivity to home observation: A comparison of audio recorded behavior with observes present or absent. *Journal of Applied Behavior Analysis*, *8*, 181–185.

Johnston, J. M., & Pennypacker, H. S. (1980). *Strategies and tactics of behavioral research* (2nd ed.). Hillsdale, NJ: Erlbaum.

Noell, G. H., DuHon, G. J., Gatti, S. L., & Connell, J. E. (2002). Consultation, follow-up, and implementation of behavior management interventions in general education. *School Psychology Review*, *31*, 217–234.

Noell, G. H., Witt, J. C., LaFleur, L. H., Mortenson, B. P., Ranier, D. D., & LeVelle, J. (2000). Increasing intervention implementationin general education following consultation: A comparison of two follow-up strategies. *Journal of Applied Behavior Analysis*, *33*, 271–284.

Noell, G. H., Witt, J. C., Slider, N. J., Connell, J. E., Gatti, S. L., Williams, K. L., et al. (2005). Treatment implementation following behavioral consultation in schools: A comparison of three follow-up strategies. *School Psychology Review*, *34*, 87–106.

Received February 7, 2007
Final acceptance June 6, 2007
Action Editor, Ken Silverman